Appendix A

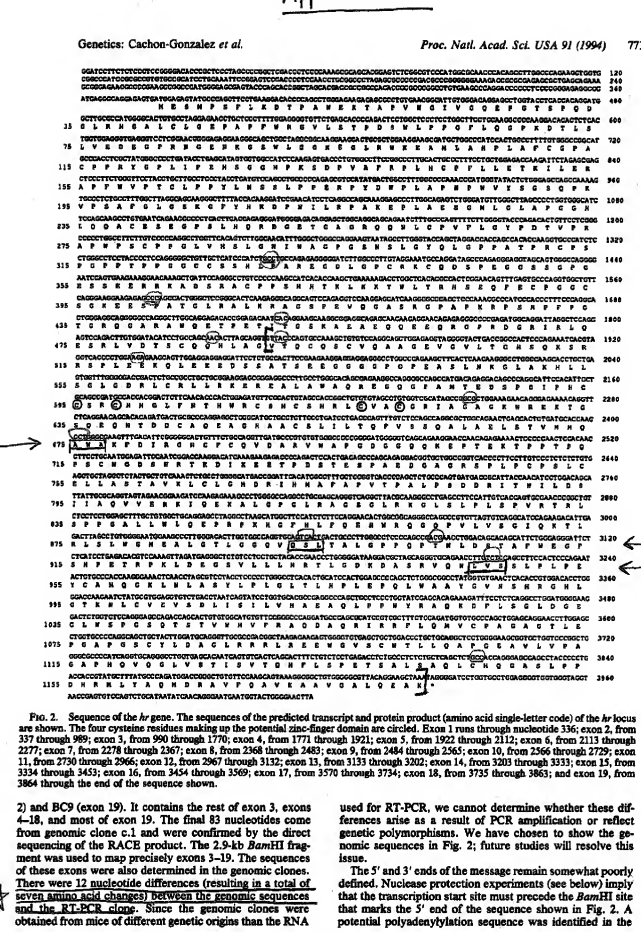

FIG. 2. Sequence of the *hr* gene. The sequences of the predicted transcript and protein product (amino acid single-letter code) of the *hr* locus are shown. The four cysteine residues making up the potential zinc-finger domain are circled. Exon 1 runs through nucleotide 336; exon 2, from 337 through 989; exon 3, from 990 through 1770; exon 4, from 1771 through 1921; exon 5, from 1922 through 2112; exon 6, from 2113 through 2277; exon 7, from 2278 through 2367; exon 8, from 2368 through 2483; exon 9, from 2484 through 2563; exon 10, from 2566 through 2729; exon 11, from 2730 through 2966; exon 12, from 2967 through 3132; exon 13, from 3133 through 3202; exon 14, from 3203 through 3333; exon 15, from 3334 through 3453; exon 16, from 3454 through 3569; exon 17, from 3570 through 3734; exon 18, from 3735 through 3863; and exon 19, from 3864 through the end of the sequence shown.

2) and BC9 (exon 19). It contains the rest of exon 3, exons 4–18, and most of exon 19. The final 83 nucleotides come from genomic clone c.1 and were confirmed by the direct sequencing of the RACE product. The 2.9-kb *Bam*HI fragment was used to map precisely exons 3–19. The sequences of these exons were also determined in the genomic clones. There were 12 nucleotide differences (resulting in a total of seven amino acid changes) between the genomic sequences and the RT-PCR clone. Since the genomic clones were obtained from mice of different genetic origins than the RNA used for RT-PCR, we cannot determine whether these differences arise as a result of PCR amplification or reflect genetic polymorphisms. We have chosen to show the genomic sequences in Fig. 2; future studies will resolve this issue.

The 5′ and 3′ ends of the message remain somewhat poorly defined. Nuclease protection experiments (see below) imply that the transcription start site must precede the *Bam*HI site that marks the 5′ end of the sequence shown in Fig. 2. A potential polyadenylylation sequence was identified in the